DOCUMENT RESUME

ED 349 827 FL 020 646

AUTHOR	Balu, Raj
TITLE	LEA Title VII Program Evaluations. Panel Presentations.
PUB DATE	Aug 92
NOTE	26p.; In: Focus on Evaluation and Measurement. Volumes 1 and 2. Proceedings of the National Research Symposium on Limited English Proficient Student Issues (2nd, Washington, DC, September 4-6, 1991); see FL 020 630.
PUB TYPE	Viewpoints (Opinion/Position Papers, Essays, etc.) (120)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Bilingual Education Programs; Elementary Secondary Education; *Limited English Speaking; Program Descriptions; *Program Evaluation
IDENTIFIERS	*Elementary Secondary Education Act Title VII; Los Angeles Unified School District CA

ABSTRACT

These panel presentations focus on LEA Title VII Program Evaluations. Raj Balu, an administrator of bilingual programs in Chicago presents information regarding the bilingual education program in the Chicago public schools, as well as information on Title VII programs and what kind of evaluation is being done. Jesus Salazar, who is currently evaluating the Eastman Project for the Los Angeles Unified School District (LAUSD) provides background information on the LAUSD, describes the Eastman project as an example of an effective academic excellence program, and offers recommendations for Title VII evaluations. Tomi D. Berney of the New York City Public Schools gives a detailed description of the process of evaluation and suggests that in order to maximize program effectiveness, it is of utmost importance to prioritize evaluation. In doing so, it is essential to continually assess both the process and the product of evaluation and to modify them as necessary. A discussion of the panel presentations by Robert Martinez of the University of New Mexico is appended. (VWL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

LEA Title VII Program Evaluations: Panel Presentations

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Raj Balu
Chicago Public Schools

I'm going to present my point of view from the perspective of an administrator. I was an evaluator, but I have been removed from that for the past five or six years. I've been an administrator of bilingual programs in the Chicago public schools a few years. First, I will be presenting information regarding the bilingual education program in the Chicago public schools. Then, I will move on to Title VII programs specifically and what kind of evaluation we are doing. From there, I will talk about the problems we have faced in evaluating the program and some strategies and recommendations. That's the basic outline of my presentation.

Chicago has about 45,000 LEP students, and the state mandates that bilingual education be provided for every child who has been identified as a LEP student. As part of meeting the mandates of the requirements, Chicago public schools assess every child who enters the school system, including the English monolingual and English background children, using a home language survey as to their language background. If any student is identified as coming from a home where a language other than English is spoken, then that child is further assessed in terms of his or her English language proficiency.

Each school has a computer terminal and the schools -- the staff in the schools -- enter the information online, and this information is available online in a central computer system. At the time the data are collected, the student is categorized as knowing no English at all, a little bit of English, or a lot of English but still needs some assistance, or is capable of functioning in a classroom where English is the only language of instruction. For any of the first three, children are categorized as limited English proficient children, and they are provided with bilingual education programs. At this time, we have about 320 schools providing bilingual education programs in about 74 different languages and about 1500 state certified bilingual teachers. Bilingual or ESL-certified teachers provide services to these children. A total of about $36 million is spent by the local schools, about $28 million by the state, and about $2.5 million from the federal government's different programs. That is the range of the program's expenditures.

I have a handout that provides these kinds of statistics, but it is limited in quantity. If you don't get the handout, please provide your

ED349827
FL 020 646
BEST COPY AVAILABLE

name and we'll mail it later. So the online data keeps track of the student's progress, and annually, at the end of the year, during the spring -- March, April, May -- each student is assessed using a citywide program. Those data are also sent to the school and, during the summer, the child's English language proficiency is reassessed, and the student is recategorized if he or she shows additional English language proficiency and is ready to move out of the program. Those not ready to move on stay in the program. This is in terms of the city of Chicago, following the state mandate. As you might have heard before, Chicago is going through the school reform initiative that was passed through the state legislature, that is, each school has a local school council, and the local school council has the right to implement the program that it wants implemented in its schools. Of course, mandated programs, such as bilingual programs, have to be provided in that school.

We have a research evaluation department, and we have a language and cultural education department. The difference between the two is that research and evaluation is in charge of managing the data, providing the evaluation reports and research reports, and planning current studies. The language and cultural education department administers the program and provides technical support to teachers in planning and implementing the programs. The two work in tandem serving the bilingual education program and providing an annual evaluation report to the state within about a six-month period after the program period. Last year we had two-fifth year programs, Title VII programs, and two new programs. One of the new programs is for Arabic bilingual education, and the other one is a developmental bilingual education program.

The developmental bilingual education program has about six schools with specific programs. All of them are Spanish-English developmental programs, and they started functioning as a program during the middle of last year, starting with kindergarten grade level. The plan is to follow the children from kindergarten through the eighth grade, even after the Title VII funding ceases. When we received funding for the developmental bilingual program, we came for a management institute, and a packet of evaluation forms and data collection forms was given to us. Those sets of forms were very useful in planning the evaluation actively for those programs. Later I will explain why they were helpful.

In the Arabic bilingual program, what we have now is that same data that we collected for the state bilingual program, the online data. Annually we are required to provide evaluation reports to the federal government for Title VII bilingual programs. I am sorry to say, we have been consistently late. I don't know whether anyone else has been late, but we have been late, and there are reasons for that. At the same time, I can assure you that the developmental bi-

lingual program evaluations report will not be late, because there was a system set up for that. It asks what things are not being done now. What are the reasons why we are not doing things that we are supposed to do? Or, what are the reasons why we would like to improve upon what we are doing in other departments?

Just within the last two months, the research evaluation department sent home 50 of its 75 staff members. The reason? Budgetary reduction, administrative staff cuts, and a perception in the schools that the administrative structure is too big. Out of the 50 people who were sent home, three were bilingual evaluators, paid out of state bilingual monies. The system is not saving any money, but in order to reduce the perception that the bureaucracy is smaller, the three were sent out. What will happen to the evaluation report that is due within the next year? It is going to be a little bit late.

Another problem we are facing is that local commitment for evaluation and research is not as big as it was in the past. Now the commitment is toward the community to improve the school's program and to show that the test grade equivalent score, even now that's being used in Chicago, has improved a lot, and the students have reached the national norm as required by the state legislature. And as the commitment is for doing the citywide testing at the end of March, or April, or May, and show that the test score has improved. Further, this assessment is required for the city regarding the general program of instruction, not just bilingual education. All of our students are part of that program, part of that assessment. So what happens? The priority for writing that report goes to the citywide testing data report and bilingual education's report stays back.

The third problem or concern in having an evaluation done is about local autonomy, the local school council I mentioned earlier. The council has the autonomy to decide what test it administers, how it administers, how it uses the data, and how it reports the data. Still, we maintain a little control of that because of the state mandate, otherwise we would have lost that control also. I'm not saying that we should control but, to collect the data that are uniform and usable citywide, control is needed.

Under the change in administration the principals are under the local school councils. They are hired/fired by them. The councils are for two years, the principals' contracts are for three years, and when a new council comes the old principal is fired and a lot of principal changes are happening. Increased costs for consultants for evaluating now average $200 per day in Chicago and non-consistent availability of external evaluators and writers. We need writers, but at the same time we hired the students from the local university to write the evaluation report, to do the evaluation, but they are not available all the time.

What we are doing now to overcome these problems and to improve upon the system is to pool the resources. Most of the Title VII evaluation that we are now doing uses part of the state money also and, as such, the state evaluation supplements the evaluation of the Title VII program. We are using college and undergraduate students as testers, and they are sent to schools to provide additionally needed testing. We are assigning staff members to be in charge of the programs, those who have an interest in a particular program. For example, our developmental program manager, there is no specific manager funded by the program, but we have assigned somebody to be in charge of the program, in addition to other duties, who has a personal interest in the program and will do a better evaluating task. In terms of Title VII guidelines, I am going to discuss the three major categories.

We collect student data diligently. We have it online, and it is available for us to analyze and study. Technical standards are maintained in terms of selection, administration, and training. We do not collect as much implementation data as we would like to collect because it is staff intensive. In the developmental program package that was given to us, we have shared that responsibility among the staff of the program in the schools because there are specific forms to be filled out by teachers and the principal of the school with regard to the data that is needed and that is supposed to be collected. In general, we collect implementation data and run it through the state bilingual programs' compliance review, which is done for a third of the schools. So a third of the schools get it without a problem. Other schools have a problem.

Lastly, I would like to talk about a few recommendations. Just like the developmental program package that contained specific data collection blank forms, which is similar to the abstract that was presented earlier by Tomi, it's a little bit more lengthy in detail, and it is divided up into different people's responsibilities to fill out those forms. This would be good for all programs. In order to collect implementation data, it may be good to have support from the evaluation assessment centers instead of the local school collecting those data. If possible, evaluation assessment centers have funds to have local persons to come and collect those data and link the data that is collected through the forms with the implementation data so that additional research can be done. But locally, if we have to do that in the collection of implementation data, we need additional money, additional resources.

In conclusion, we had an agenda last year and tabled it. We wanted to do a longitudinal study of bilingual program students to present those data to the state legislature. We tabled it because we did not have staff, but we have plans to go back and do that within the next two or three years if everything comes out right.

Jesus Salazar
Los Angeles Unified School District, California

I am going to talk about the Russian revolution, an on-going statistical revolution, and Title VII programs. You'll see how these three themes are related as I make some recommendations for Title VII programs. I'm currently evaluating the Eastman Project for the Los Angeles Unified School District (LAUSD). It is a seven-year longitudinal evaluation study that follows limited-English proficient students from kindergarten through the sixth grade. Let me give you some background on LAUSD. It has roughly a quarter of million limited-English proficient students. LAUSD's LEP population alone would make it the fifth largest school district behind only New York City, Los Angeles, Chicago, and Houston.

I am not going to explain the Eastman Project curriculum, that's another story. Suffice it to say that the Eastman Project served as the basis of LAUSD's Bilingual Master Plan that was implemented districtwide in 1989. Prior to the Bilingual Master Plan, primary-language instruction was provided by para-professionals in more than half of the bilingual classrooms. Schools implementing Master Plan models similar to the Eastman Project could reduce the number of bilingual classrooms by as much as 33 percent.

The seven-year longitudinal evaluation study I am conducting has resulted in a Title VII Exemplary Academic Excellence Award. At the time I began this evaluation study I did not know that I was going to be conducting Title VII Grant research. I've been learning as I go. The 1991 Russian Revolution: Paradigm for Statistics.

There is a major revolution going on in Russia as I speak. Who would have imagined the radical changes now occurring after more than 70 years of communist rule. I use the second Russian Revolution as an analogy because a similar revolution is occurring in statistics. The revolution in statistics began in the 1960s and is occurring at a slower pace than the Russian Revolution, but it is a revolution nevertheless. Back in the 1920s a major political-paradigmatic debate took place among statisticians in the social sciences and applied statistics. The "party" that finally won decided to report statistical findings in terms of levels of significance.

For those of you conducting evaluation studies and applied research in educational settings, you know that most parents, teachers, and school administrators couldn't care less about the probability level of a study. Quite frankly, I find research that reports only levels of statistical significance very boring. Ultimately, this type of research does not tell you very much. Parents and the educational

community want to know only two things: first, ***does the program work?***, and second, ***How effective is the program?*** Unfortunately, the "party" that lost the statistical wars in the 1920s was in favor of reporting statistical results in terms of program effectiveness. That is, rather than reporting that a study was statistically significant at the .05 level with 34 degrees of freedom, these statisticians were more concerned in showing that variable X was 86 percent more effective than variable Y in improving reading scores. Any parent or teacher will relate to a study that shows a program can help a child learn to read 86 percent better.

Title VII Grants and Measures of Program Effectiveness

The California Department of Education has two requirements for Title VII Grant research and evaluation applications. First, the traditional level of statistical significance needs to be reported. Second, program effectiveness relative to a comparison program has to also be reported. The State Department of Education essentially has you do a "Pepsi Challenge Test." I was very happy when I heard Ms. Sevilla discuss her major concern about integrating the ivory tower research community with the public community. I believe that reporting data in terms of the effectiveness of a program begins to address her concern. That is, the research community can continue to conduct multi-variate analyses with sophisticated research designs, yet the findings can be reported in terms of program effectiveness. We can all benefit from this type of research paradigm.

I conducted a three-year longitudinal study and performed multiple analysis of variance (MANOVA) statistics, yet I have been able to make this data meaningful to parents, teachers, and administrators by reporting the academic effectiveness of the Eastman Project. This approach to data analysis is so applied oriented that I have been able to present the positive longitudinal effects of bilingual instruction to parents in Spanish.

Two questions are always asked of me whenever I do presentations before parents and teachers. First, parents want to know, ***"Is my child learning English?"*** That's what they basically want to know. Second, teachers ask me, "Is the program I'm teaching in working? How effective is the program and why should I continue using this instructional model instead of the model that I was using?" These questions were common both in the initial phase of the Eastman Project implementation and in the initial period of the Bilingual Master Plan implementation.

EFFECTIVENESS OF EASTMAN PROJECT COMPARED WITH COMPARISON SCHOOLS AND DISTRICT

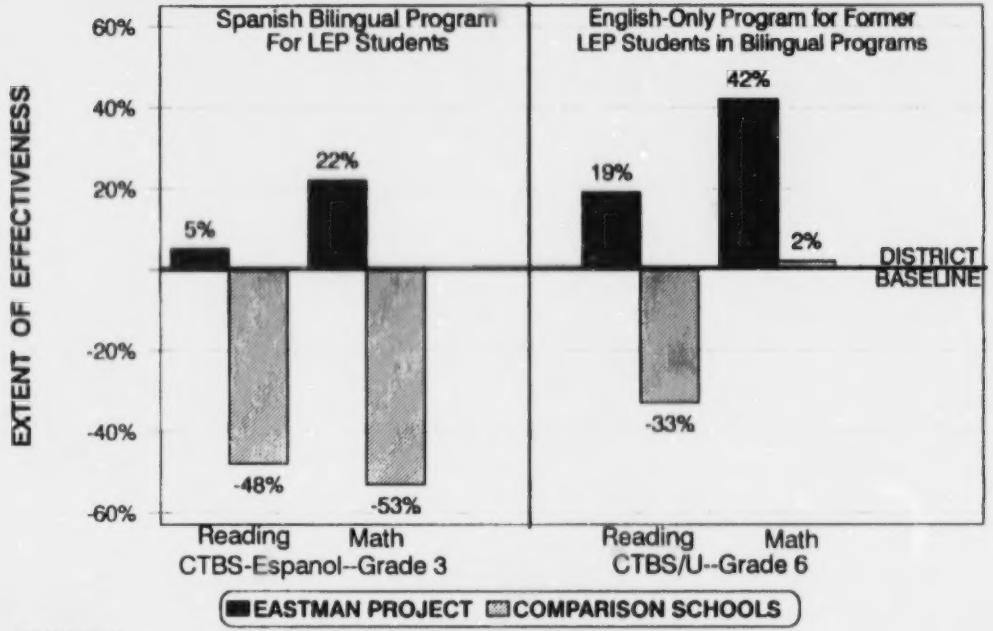

FIGURE 1

Third-Year Implementation (1988-89) CTBS Scores

Eastman Project:
An Effective Academic Excellence Program

After three years of implementation, third grade limited English proficient students who received three years of Spanish-language instruction in the Eastman Project were learning to read in Spanish 5 percent better than LEP students districtwide and 53 percent better than a group of LEP students who received three years of instruction in a comparison bilingual program (see Figure 1). LEP students at the Eastman Project schools were also learning Spanish-language math 22 percent better than the district baseline and 75 percent better than the LEP comparison group. In short, the Eastman Project is a more ***effective program*** for teaching LEP students in the Los Angeles Unified School District.

Though these findings of Spanish-language instruction are very encouraging, the key question for educators and parents is, ***how well does this knowledge acquired in the Spanish classroom transfer into English-only instruction?*** We have a preliminary answer, but a very encouraging one. Figure 1 also shows that the English Project has been successful in providing English-only instruction to students who previously received Spanish-language instruction. Briefly, the Eastman Project was 19 percent more effective in teaching English reading to former LEP students than the District English reading baseline. The project is also 42 percent more effective in teaching former LEP students to read in English than the comparison school English-only program. The Eastman Project was also 42 percent more effective than the District in teaching math in English-only classrooms, and 44 percent more effective than the comparison school program.

As I mentioned earlier, what makes these results exciting is that not only were these ***statistically significant findings*** but that the results were also presented in terms of program effectiveness. That is, every parent I've met would prefer that her/his child be enrolled in a program that is 19 percent more effective in teaching the child to read English. One of the recommendations that I have for Title VII grant applications and reports is that program effectiveness be presented graphically as I have in this presentation. This is the type of graph that you see in the Wall Street Journal. This data, as I indicated earlier, is based on a MANOVA analysis. Yet, when presented graphically and in terms of program effectiveness, the data become even more powerful. Again, let me emphasize the practical applications of this model. The findings of this model can be used in parent-teacher conferences to provide parents with information regarding the most effective programs for ultimately teaching English to LEP children. In the Los Angeles Unified School District, parents can choose whether or not to enroll their children in bilingual programs.

This information about program effectiveness can be provided to parents to facilitate their decision.

Recommendations for Title VII Evaluations

I want to identify four models for evaluating Title VII programs. These models are based on my search of the research and evaluation literature over the past two years. These models are listed in order of importance for reporting the effectiveness of instructional programs. That is, evaluation models should first and foremost highlight the effectiveness of an instructional program. If statistical significance needs to be sacrificed for the sake of evaluating program effectiveness, then so be it. The four models are listed below in their order of importance for Title VII evaluations:

1. A program is both educationally and statistically significant

2. A program is educationally significant and statistically non-significant

3. A program is statistically significant and educationally non-significant

4. A program is both educationally and statistically non-significant

A program is considered to be educationally significant if it is demonstrated to be a more effective instructional program when compared to another program. Studies exist where highly statistically significant findings were obtained but yet were basically educationally non-significant (e.g., one program was 3 percent more effective than another program). There have also been instances where an evaluation study did not reach statistical significance, yet one program was demonstrated to be 15 percent more effective than another program in teaching students to read English. Under the current research and evaluation paradigm, the case where statistical significance and educational non-significance is obtained is considered more noteworthy than the case where statistical non-significance and educational significance is demonstrated. However, for practical applications the latter case carries greater educational importance. As I mentioned earlier, most parents would rather have their child to be in a program that is 15 percent more effective in teaching English reading than a program that is only 3 percent more effective. Statistical significance be damned!!! After all, statistical significance is not able to teach Johnny or Juanito to become a better English reader!

Epilogue

Statisticians and researchers in the areas of meta-analysis and power analysis are leading the change in re-emphasizing the effect size of programs. Jacob Cohen, who is my statistical hero and the preeminent statistician today, was the first to emphasize effect sizes in the 1960s. Thirty years later statisticians in the social sciences and educational research are still reporting the majority of their findings in terms of statistical significance. However, with others such as the prominent Harvard psychologist Robert Rosenthal leading the way, measures of program effectiveness have become more common in the 1980s and 1990s. The revolution to report research findings in terms of program effectiveness is thus gathering momentum.

I want to close with one of my favorite quotes attributed to Henry Ford. Ford is quoted as saying that...***"if you can't write down your idea on the back of my business card, then you don't have a good idea."*** This reflects my sentiments regarding program evaluation and research. That is, many researchers do not present their evaluation findings concisely and to the point. Let me put it this way, which idea would Henry Ford have been most likely to be impressed with, to write on the back of his business card that the Eastman Project is 42 percent more effective in teaching English reading to former LEP students; or that the difference between the Eastman Project and comparison school reading programs is statistically significant at the .05 level as analyzed with a MANOVA with 75 degrees of freedom using a repeated measures design...?

Tomi Deutshc Berney
New York City Public Schools

Title VII evaluators are intermediaries. They stand between the providers of information and its consumers.

First, the providers of information: Project staff know they must report a great deal of information. While we understand that project directors frequently have an inordinate number of tasks to accomplish, there is information that only they can provide, and providing accurate information is a critical responsibility.

Understandably, projects are loath to transcribe data which they've already supplied to someone else; they resent searching for numbers that, it later turns out, no one needs to know; they are opposed to giving students unnecessary tests. Neither project personnel nor evaluators wish to place unnecessary burdens on students and teachers. No one should be asked to perform superfluous tasks. It is the obligation of the evaluator to make the evaluation process as efficient as possible. We, in New York City, are currently seeking to do this.

Second, the consumers of information: Evaluation reports are not found on paperback book racks; they are not read for pleasure. Consumers want to inspect the instructional and non-instructional data provided on past and current program participants. Consumers want to study the information about program activities and materials. They want to know what the impact of the program has been on student achievement in English or the native language where appropriate, and content area subjects. They want to discover whether the project has met its specific program objectives. The consumers of information justifiably expect the evaluator to assist them in accomplishing these tasks as quickly and efficiently as possible; they value clarity and conciseness.

The process and the product of Title VII evaluations are equally important. In both, we're learning as we're doing. We have not yet finalized either how we are going about collecting information or in what way we are going to report that information.

The process of evaluation should be as uncomplicated and efficient as possible; the product of evaluation should be clear and concise. In enuring that both the process and product are as they should be, the evaluator must perform a myriad of tasks: prepare forms which are easy to complete accurately and fully, collect the data and analyze it, integrate information from a wide variety of sources, and write and edit the report. Evaluators must go through these steps each project year.

Despite this, implementing a process that is uncomplicated and efficient and providing a product that is clear and concise are possible. The Office of Research, Evaluation, and Assessment (OREA) of the New York City Public Schools has developed a plan to implement the process of efficient data collection and has a prototype for a succinct, information-filled report. But neither the process nor the product is static; an effective evaluation system should be adaptable.

The process of gathering information is based upon two principles: First, avoid asking for information unless you know exactly how you will use it. And second, inside almost every open-ended question is a concealed closed-ended question. When possible, ask the closed-ended question.

Technology is the means which offers the greatest opportunity to streamline the process of collecting data. Because electronic records can be accessed and updated so much more easily than can paper records, electronic record-keeping must be the medium of choice. In New York City, we have taken steps to ease the burden of reporting data by utilizing an electronic system wherever possible.

A central computer maintains all citywide test scores; many offices can access these scores. Title VII programs, therefore, need not report their students' scores on the Language Assessment Battery, the instrument with which we measure proficiency in both English and Spanish, or on any citywide tests. Once we have a participating student's name and student identification number, we can automatically get pre- and post-test scores.

We ask project staff to complete a Student Data Form for each participating and formerly participating (now mainstreamed) student. Most of the background information required by Title VII comes to us on this form. When we receive the Student Data Forms, we enter the information they contain into the system. The following year, we preprint the Student Data Forms for continuing and mainstreamed students, using this background information already in the system. Project staff need only report any new information, including data on attendance and academic performance.

We used to ask that separate forms be completed in the fall and in the spring for high school programs. This year we developed a single form with fields for data for both semesters. Ultimately we would like to reduce the burden on project staff even further by retrieving individual attendance rates and course grades from the central computer files.

To learn about staff qualifications and program activities and materials, we developed a Project Director's Questionnaire (P.D.Q.). This is followed by a structured interview. We have tried to make

both the P.D.Q. and the interview as efficient as possible. We ask for information that Title VII requires. We also ask for explanatory information to help the consumer of the information better visualize and understand the project.

The product, the evaluation report, is obviously at least as important as the process of evaluation. The report should be clear, concise, and filled with information. Title VII regulations may dictate the content of the report, but the evaluator can choose the form. In New York City we have developed what we call a profile format. In developing the format, we first worked from a list of Title VII regulations and made sure that each regulation was covered in the report. We then held focus groups and spoke to project directors and personnel from the State Education Department of New York. They recommended changes that would make the reports more useful to them. These recommendations were extremely helpful.

The profile format has gone through a number of changes. It's still going through changes. We are learning as we're doing. For an evaluation system to be efficient and effective, it must be adaptable. It can't be static because your populations change and your priorities change. It is very important that you change the evaluation report to meet the changing needs.

The report has two parts: the Extract and the Program Assessment. The Extract is not a summary. It presents the salient points, the most important points, the items which we were asked to state on the first page. Since there is too much information to state on one page, this section of the report is contained in the first two pages. The Extract contains information on the funding cycle, sites, enrollment, background of the students served, admission criteria, programming features, and strengths and limitations. It also gives OREA's conclusions, including which objectives were met and which were not, and recommendations.

The funding cycle indicates what year of funding the project has just completed. The sites section lists the sites in the project, the grade levels included, the number of students participating in the program at each of the sites. Student background lists the number of students by native language and country of origin. We include here information on how many years of education the students had, on the average, in their native countries, how many they had in the United States, and what proportion of students were eligible to participate in the federally-funded free lunch program. Admissions criteria includes any criteria the program uses for program participation.

Programming features, strengths, and limitations of the project presents those and states what objects the project met, which it did

not meet, and for which it provided no information. Reasons why a project may not have achieved a certain objective are given later. In the recommendations section we frequently recommend exploring reasons why objectives were not met, suggest that objectives be modified to make them more realistic, or suggest ways of meeting objectives or providing data. We try very hard not to recommend things that necessitate the expenditure of additional funds.

The Program Assessment is the major part of the report. Its sections are: staffing, implementation and outcomes, services to students with special academic needs, mainstreaming information, and a case history.

Staffing lists the title, highest educational degree, and language competencies of the Title VII-funded staff. Other staff who work with project students (teachers, for example) are described in aggregate.

The second section, implementation and outcomes, is structured around the objectives of the project. For each objective, we report relevant activities, the evaluation indicator used, and a summary statement as to whether the project either met or did not meet the objective. The section may include teaching techniques and materials, or the latter may be listed in an appendix. Whether or not there are objectives concerning attendance and dropout rates, information on those is presented in this section.

The part of the report where objectives are presented and discussed is critical. Objectives define the direction of a project. A well-stated objective helps the project tell the world how good it is. A well-written objective clearly states who is expected to accomplish it, what the expected performance is, and when the accomplishment of the objective will occur. Unless objectives fulfill these criteria, they should be considered unacceptable.

The third section shows statistics on students with special academic needs. This includes data on students referred to special education, to remedial programs, to programs for the gifted and talented, and how many students were retained in grade. We include here the linguistic competencies of the school staff who evaluate students for these programs. We look both at the number and the percentage and we attempt to compare the current year's data with those of the previous year to see if there has been any change. Has the project made a difference?

The fourth section gathers information on students mainstreamed and the number of graduating students planning to enroll in postsecondary education institutions. This section concludes with a report on the academic progress of former project participants who have been mainstreamed.

The case history, though not required by federal regulations, is the fifth section. It suggests the program's impact on individual students. The case history gives the consumers of the information the ability to visualize and to understand the project that we are evaluating. In the case history, we carefully maintain confidentiality.

We feel that the Profile Format will meet the needs of those involved in Title VII projects as well as state and federal officials who review the achievements of these funded programs. With further experience and continued feedback, we will continue to refine the form. We have added a page to the profile format that explains how we gather the various kinds of data as well as the statistical procedures we have chosen to use.

One additional advantage of a clear and focused format is that it facilitates the preparation of academic excellence applications. The profile format provides information in a way which simplifies the task. It can be easily determined whether a project may be considered to be exemplary in any area. A description of participating students and staff, program activities, academic and non-academic achievements, and the degree to which the project met its objectives are presented clearly and concisely.

Evaluation should be a high priority. We feel that this priority should be reflected in the number of points allotted to evaluation on the Title VII grant application, and more specifically, that it be required that objectives be clearly stated, measurable, and realistic. An evaluator can easily assist a proposal writer or a prospective project director develop well-written objectives. The New York City Public Schools Office of Research, Evaluation, and Assessment routinely offered this service. Those who write proposals and those who approve them should place greater emphasis on objectives and project evaluation. It is hoped that the reauthorization of Title VII for 1993 will address this issue.

Three things are necessary to improve the quality and value of Title VII evaluations. First, the process of evaluation should be as uncomplicated and as efficient as possible; this is becoming more attainable with electronic record keeping. Second, the product should be informative, clear, and concise; these are the goals of the Profile Format. Finally, objectives should be well-formulated, clearly stated, measurable, and realistic.

In order to maximize program effectiveness, it is of the utmost importance to prioritize evaluation. In doing so, it is essential that we continually assess both the process and the product of evaluation and modify them as necessary.

Discussion of Panelists
Balu, Salazar, and Berney's Presentations

Robert Martinez
University of New Mexico

What's being passed out are the Title VII regulations with which all LEAs receiving Part A funds are required to comply. After they are distributed, I would like to talk about them with you and the group. This is basically what Alan Ginsburg had called "the laundry list" on his first day presentation. This has been our laundry list, so-called, for the last six years. The reauthorization of Title VII will be coming up very shortly.

It is now time to do, in testing terms, an item analysis with these Title VII evaluation regulations. We need to look at those regulation items -- those we need to keep, those we need to revise and, most definitely, those we need to exclude.

What I'd like to do at this point is to address six specific regulations since we don't have time to address them all, but six that I feel have credence. They had a lot of problems in being addressed or not being addressed in the field. I would like to address them by first stating the regulation and then having my colleagues address how they work with that regulation and their respective school district.

First one. Under 500.50(b)(1), a grantee's evaluation design must include a measure of the educational progress of project participants when measured against an appropriate non-project comparison group.

Before my colleagues respond to this, I would like to say that we've looked at these evaluation requirements in addition to regulation requirements required by Chapter 1, Special Education, Indian Education, Migrant Education, and we find that these are the most comprehensive and the most stringent. However, they do provide for program improvement which our programs are all about. With that, I'd like to work eastward starting from the west.

Jesus, would you mind addressing the requirement that a grantee's evaluation design must include a measure of the educational progress of project participants when measured against an appropriate non-project comparison group? Could we just limit our discussion to maybe a minute for individuals?

Jesus Salazar

You caught me off guard, but I can try and answer it. I think I was mentioning in my presentation that we actually have two kinds of comparison groups with district baselines, but we also have a group of schools very similar in demographics to the project schools. We just use traditional California Test of Basic Skills test scores, California Assessment Test Scores, but the extra little twist that I add, and that's the one I mentioned, is effect sizes. I also report the effect sizes for the comparison schools that must report.

Robert Martinez

Is there a certain design that you use most often?

Jesus Salazar

I have conversion tables for anyone who is interested. Any statistics test, any research design you use, after you've run the statistical analyses, you can convert that analyses in effect by a multiple analysis of variance (MANOVA), one way, two way, two by four, you name it, and I can provide an effect size.

Raj Balu

We have in Chicago different kinds of comparison groups identified for Title VII programs. One is within a school, and one set of comparison groups within school groups as well as between school groups. We have Title VII programs and, within the same school, we have transitional bilingual education programs mandated by the state for students who are not receiving Title VII services. So those two groups can be compared, and we also have Chapter 1 students and English monolinguals, English students who receive English instruction only, not even transitional bilingual education services from the state. So there are four different groups that are used in our data analyses: Title VII group; the state bilingual program without Title VII; students with no bilingual program, that is English only instruction; and, finally, Chapter 1 students who are bilingual and receiving services under Chapter 1 and bilingual services. They are receiving two sets of services very similar to Title VII.

Let me qualify one more item. The students who are in Title VII programs receive two kinds of services -- the state bilingual services and the supplementary services from the Title VII programs. Under the developmental program, we are thinking of using the grade-design and follow them for up to eight years; that is, we follow different students who get into the program at different times.

Tomi Berney

In New York, we really have a different set of circumstances. In New York City, every student who is limited English proficient is entitled to receive supplementary services, so we really can't do a control group design, an experimental group/control group design. What we use is a gap reduction design and, instead of using an equivalent control group, we use the norming sample, the group on which the language assessment battery was normed.

Those of you who have our sample report, on the bottom of page 3, we've specifically discussed this, how we use the gap reduction and the way in which we use it is we talk about normal curve equivalence (NCE). When you're talking about NCEs, you are assuming that there will be no gain, there will be zero change in NCE from year to year; we use Spring-to-Spring evaluation. When there is a gain, it means that the participating students are doing better than we would have expected the group on which the test was normed to have done. Assumedly, they will do better than just one NCE better, we would hope for at least five but, in any case, it is a gap reduction design that we are using.

Robert Martinez

The second regulation I will address is 500.50(b)(2)(ii), reliability and validity of the evaluation instruments and procedures. The evaluation instruments used must consistently and accurately measure progress toward accomplishing the objectives of the project, and they must be appropriate considering factors such as the age, grade, language, degree of language fluency, and background of the person served by the project. I'm particularly interested in addressing those populations where standardized tests are not available. What do you use at that point?

Raj Balu

First, let me answer by explaining the kind of standardized test that we are using now in the Chicago public schools, the first one initially that we used for students admitted to the school programs. It is a functional language assessment instrument. It's a simple instrument that is testing the auto language skills of the students, and we have questions about the validity of that instrument, and we are in the process of revising that.

The second group of instruments is used during the initial enrollment of the students into the bilingual program; this is their language assessment scale, which is a standardized test; it is being used now. One of the concerns we have is that it is time-consuming, mainly the auto component of that particular test is time-consuming.

The reading and writing we are going to continue to use for some time.

During the spring, we are using the Iowa Test of Basic Skills for the first grade through the eighth grade level and then the Test of Academic Progress in the high school grades. These are the same tests that are used for citywide testing, so that we will have comparable data for viable programs and students who are in different kinds of program situations.

When tests are not available, an example of a situation I can say right now is that the auto language component assessment has become a little difficult for us. The language assessment scale takes about 45 minutes for each child. We have 45,000 students, and we are required to test every one to assist the progress of the students. Currently what is being done is that the teacher is asked, read to the children; develops oral language proficiency from one through five. That's all that's being done, and we are trying to bring into the system the ways that they rate the children and ways of increasing the validity and objectivity of this particular process.

Tomi Berney

The language assessment battery that we use in New York is reliable, and it's valid. In fact, it was just renormed this current year. It has four components -- listening, speaking, reading, and writing. When students come into the system and are initially tested, let's say in the fall or in the middle of the year, they take what's called a short version of it. This does not include reading, and it doesn't include speaking. But on this score we determine whether a student is limited English proficient or not.

Then when the spring testing takes place, we give the full lab which again really does not include the speaking sub-test, but that's something else. It includes the listening, the reading, and the writing components, and we can measure from year to year how the student is doing. We do use the speaking component for something else; it's not that we ignore it; we just look at that score separately.

The norms that we use, the lab itself, are the English proficient norms. We're not using limited English proficient norms. A student, to show growth, really does have to gain in skills. However, it is reliable, it is valid. In some cases, we do use other citywide tests depending on whether the student has been in an English speaking school system for at least two years. Where we run into problems is testing students in languages other than English. In Spanish, we have no problem because the lab is in Spanish also. We are in the process right now of developing a Chinese reading test, which we really don't have at this point. The other major language group in New

York is Haitian Creole, and we don't have a test there. What they have been using in those cases are either teacher developed or district developed tests.

Jesus Salazar

There is a section in Los Angeles, actually it's the Hollywood section, which is better known for the stars' walk of fame, etcetera, that has 60 languages represented. So you can imagine the situation teachers in schools are facing. We rely to a large extent on the Student Oral Language Observation Measure (SOLOM), which is teacher assessment. It takes about 20 minutes, and teachers usually give it about a month after they have had a student in their classrooms. That's the measure we use to identify a student's level of English proficiency or lack thereof.

We are also in the process of developing an Asian language curriculum for the district. We have, I believe, 40,000 limited English proficient students. We have a lot of Armenians, a lot of Russians, and overall the district has about 88 languages represented. We're in the process of trying to address as many as we can. We rely on the CPPS as a measure for transitioning students from native language instruction into English language instruction. They have to meet the 36th percentile, so we rely a lot on norm testing once again, English. Prior to that, as I mentioned, we rely on teacher observation measures to identify for placement in programs.

Robert Martinez

The first day of the conference was focused on alternative assessments for performance-based assessment. Is it now time for that to be included in Title VII regulation requirements? I'm not going to turn it over to you for an answer at this point, I'd like to continue the other ones. But it is food for thought in that the authorization will be up. You may want to consider that and address it with the appropriate personnel.

Under 500.50(b)(3)(i)(B), evaluations must provide information on the academic achievement of children who were formerly served in the project as limited English proficient, have exited from the program, and are now in English language classrooms. How has Chicago addressed this requirement?

Raj Balu

The current exit criteria for LEP students is that they need to be at the 50 percent cut off point. That is, they should perform at the fifth stanine before they are ready to exit from the program. There is a conditional exit that is the fourth stanine and that is a recommen-

dation. All of these are on the Iowa Test of Basic Skills or the Test of Academic Progress, depending on the grade level of the student. Now, once the student exits, we follow the student for a year and, if the teacher recommends this child be brought back into the program, then we bring the child back into the program. THIS IS REQUIRED BY THE STATE LAW.

Robert Martinez

What about the children who have exited from this program?

Raj Balu

The exited students, especially the conditionally exited students, are followed for at least one year and, if necessary, brought back for support services, not bilingual education but transitional programs.

Tomi Berney

In New York, I mentioned before that we distribute student data forms. One of the student data forms is for previous, now-mainstreamed program participants. On this form, the school must give information concerning class grades, sometimes test scores, such as New York State Regency Examinations, or Regents Competency Test, and attendance data. In this way, we are able to follow the student. The one problem we have is that the schools do not like to bother filling out information; these are no longer program students, and it's very hard to find people to give us this information. That's what I was saying before, to ensure that we get the information and get it into an electronic system would be much to our benefit. I'm sure we miss a lot of students. We don't get all the mainstream students.

The other problem is when a student goes from school level to school level -- from elementary to middle school to high school -- we lose track of that student. Ultimately, New York City is supposed to be on a computer system called Automate The Schools (ATS). It's not in all districts now so we can't access this information everywhere. As far as I know there are not yet any plans to do this in the high schools. Once every school is on computer, it will be very simple to get this information on any student with the name and ID number of the student. We can't get it now. We have to rely on school personnel to give us these data. We tried to make the forms as simple as possible, but we can't force them to do it. That's the problem involved with it.

Jesus Salazar

In the Los Angeles Unified School District, all of the elementary school students are on a computerized database. We have gotten to the point where even if they move to another school within the district we can follow them because of ID codes. We have half of the junior high school students on the computerized database and half of the high school students on the same system. By the end of this academic year, 91-92, we should have all students entered into the computer database. What we're doing is following former LEP students who have exited from a bilingual program, not just Eastman but any of the other eight bilingual programs we have; we can follow their academic progress until they leave the district.

I'm currently doing a follow-up study of the original Eastman elementary school -- that's what the Eastman project is based on. It started back in the 81-82 school year. I'm following the different cohort group of students. Most of them are 10th, 11th, and 12th graders currently. I'm doing a 10-year follow-up to see if the academic achievement has been sustained over a 10-year period.

Robert Martinez

Two more, if you would, and then I'll open up to questions. Under 500.51(f) Title VII grantees must collect information on the specific activities undertaken to improve pre-referral evaluation procedures and instructional programs for LEP children who may be handicapped or gifted and talented. Chicago, how do you deal with that?

Raj Balu

We have two different departments that handle the education program of these children. One is the special education department. It is the primary one in terms of identifying and following through the referrals and assessment of the children as needed. The other is bilingual education, in which the language and cultural education department gets involved. There are two tiers of the gifted program. One is across the board for all students, and LEP students are also eligible to participate in that particular program. In addition, last year, we devised a Spanish-English gifted children's program, and that begins this year. Once this becomes a practical program in terms of planning, implementation, materials-development and so on, then the idea is to expand this to other languages as well as to additional schools in the Spanish language.

Jesus Salazar

The Los Angeles Unified School District also has a referral, teacher-test situation. Unfortunately, some of the coordinators at the school sites were of the opinion that students cannot be classified as gifted until they learn English. That's unfortunate because we've seen cases -- we've caught cases like that. We now have a system where we can bypass coordinators who feel that way. We have an office of special education that identifies gifted students in all languages, Spanish, Korean, and the other 86 languages that we have in the district. Everything is maintained on a computerized database so that at any one point we can keep track of the gifted students.

Tomi Berney

Fortunately, in New York all programs are open to LEP students -- gifted, talented, or remedial programs or any others. It is specifically stated that LEP students are eligible. As far as that certification for special education, every district has a school base support team and a committee on special education. Here there is a school psychologist, an educational evaluator, and a social worker who make a recommendation as to whether a student belongs in special education. It is hoped that, in cases where the student is limited English proficient and where the parents really don't communicate in English, the person who either tests the student or speaks with the parents is able to communicate in the language of the student and/or the parents.

In addition to the fact that all gifted and talented programs are open to LEP students, we have specific programs for the gifted and talented LEP students. I evaluate at least 60 programs so we have some of everything in New York City.

Robert Martinez

Last one. 500.52(c)(5) asks grantees to report on the extent of educational progress achieved through the project measured, as appropriate, by changes in the rate of student enrollment in post-secondary education institutions. How does Chicago handle that?

Raj Balu

About two or three years back, there was a problem of LEP students who graduated from high school taking the ESL One and ESL Two courses and the basically many, many needed English courses to get admitted into the colleges and the universities. They needed to take additional courses. In the last two years we had a committee task force working with the university and we have resolved that

particular problem. There are different standards set now under the ESL for these students who graduated from high school to get into the program and we follow some of the students -- not in detail.

Tomi Berney

The question really concerns how do we follow this. What we do, we ask a question on the student data form. First we find out what grade the student is in, and we assume 12th graders are graduating from high school as long as they're not being retained in grade and they have not dropped out, of course. We then ask a specific question, does this student plan to enroll in postsecondary educational institutions and we just compare the two pieces of data and we find out what percentage of students who are graduating from high school are going to be entering postsecondary institutions.

Jesus Salazar

In the Los Angeles Unified School District, I'm with the bilingual unit of program evaluation and assessment. We have five different units. The research unit follows students who go into higher education, and they keep track of as many students as possible. What I feel we are going to be doing in the near future is to follow through on students in different types of bilingual programs within a district who have exited into English only programs and have gone on and graduated and gone into college. We're probably doing that type of longitudinal study, and I mentioned earlier that I'm doing a study, a 10-year follow-up of the original Eastman elementary school students. Some of them, the cohort group that was in the fourth grade, are currently freshmen, those that have continued in college. I believe that I'm going to be asked to do another study to follow up that group of kids, so they're probably going to have to work for me when I get back.